WPS5694

Policy Research Working Paper 5694

The Success of Infrastructure Projects in Low-Income Countries and the Role of Selectivity

Nicola Limodio

The World Bank
Development Economics Vice Presidency
Research Support Unit
June 2011

Policy Research Working Paper 5694

Abstract

This research analyzes the success of the infrastructure projects financed by the World Bank, focusing on the causal link between the quality of project implementation and its outcome. The results show that the success of infrastructure projects depends fundamentally on the quality of implementation. Although bad implementation can harm structurally solid projects, good implementation cannot make structurally weak projects successful. This leads to the conclusion that governance and selection of well-designed projects are essential for success and, in order to improve project outcomes, multilateral development banks may need to align their incentives toward this objective and invest more in governance and capacity building.

This paper is a product of the Research Support Unit, Development Economics Vice Presidency. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at nlimodio@worldbank.org

Produced by the Research Support Team

The success of infrastructure projects in low-income countries and the role of selectivity

Nicola Limodio*

World Bank

Keywords: infrastructure, public investment, project design, aid effectiveness, low-income countries, World Bank, Multilateral Development Banks.

* nlimodio@worldbank.org or nicola.limodio@gmail.com , Junior Professional Associate, Development Economics Vice Presidency, World Bank. I need to acknowledge the fundamental support of Piero Cipollone for encouragement, guidance and useful discussions, his contribution has been extraordinary. I would also like to thank for their comments, ideas and help Jean-Jacques Dethier, Chris Papageorgiou, Antonio Estache, Michael Klein, Daniela Gressani, Sonyia Caravalho, Marianne Fay, Somil Lall, Stephan Straub, Teresa Garcia-Mila, Goldman Eitan, Mehtap Akguc, Francesco Grigoli, Aart Kraay, Shiva S. Makki, Nancy T. Lim, David Rosenblatt, Gabriela Inchauste, Paolo Verme, Mattia Ricci, Gloria Pracucci and the participants to the "Workshop on Infrastructure, Corruption and Procurement" held at the Toulouse School of Economics, May 5-6 2011. All mistakes are my own. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

1 Introduction and Reasons

Infrastructure investment is a fundamental instrument to promote development and enhance capital accumulation. It reinforces network effects (Roller and Waverman 2001), regional comparative advantage (Estache and Fay 1997), education (Brenneman 2002), health (Agenor and Moreno-Dodson 2006) and other growth factors. Its importance is out of question, as the 1994 World Development Report on infrastructure shows (World Bank 1994), and every year more than 200 billion dollars are spent on this sector in developing countries, roughly a fifth of total investment.

As well as public authorities, also Multilateral Development Banks (MDBs) regard infrastructure investment as a priority and promote not only its financing, but also design and regulation. As Table 1 reports, the five largest MDBs lend more than USD 50 billion in infrastructure per year, with this single sector absorbing between 30 and 70 percent of their total lending.

However, pouring money is often insufficient to achieve results and, as this work shows, the governance of projects, their selection and implementation are the main factors on which authorities, both national and international, need to focus.

For this reason, in this paper we study the performance indicators of infrastructure projects and focus on the micro relations involved in them. Through a rigorous empirical analysis, we explain which variables are central in getting successful infrastructure projects and what actions can lead to high performance. The results of this work are mainly directed to Multilateral Development Banks, as we explore how to boost project effectiveness and what can be done toward this aim. We believe that this is an important research field, because in times of high uncertainty in economic prospects, especially for low-income countries, getting infrastructure investment right is central in cutting poverty, sparking growth and consolidating future fiscal positions.

The empirical analysis relies on a new database we assembled, which contains 1912 infrastructure projects financed by the World Bank between 1979 and 2008; these include detailed project performance evaluations on which we build this paper. Anticipating on the results, three main insights emerge from this research: 1) the quality of project implementers (borrowers) is the most important proxy for project success, regardless of most macroeconomic variables; 2) successful project implementation is positively related to the performance of the

supervising agent (the World Bank in our case) and 3) successful implementation can create an impact only in presence of structurally solid projects, therefore selectivity and project governance are the tools through which successful infrastructure investment can be achieved.

Table 1: Infrastructure spending per Multilateral Development Bank in billion dollars*

Multilateral Development Bank	Total Lending	Infrastructure Lending	% of Total
World Bank	58.70	26.41	44.9%
African Development Bank	7.52	3.91	51.9%
Asian Development Bank	11.46	7.51	65.5%
European Bank for Reconstruction and Development	7.72	2.53	32.7%
Inter American Development Bank	13.04	10.0	76.6%

Note: * for the World Bank the figures refer to its 2010 operations, refer to the Annual Report webpage; for the African Development Bank the figure refers to its 2009 operations, refer to the Annual 2010 Report, page 18; for the Asian Development Bank the figure refers to its 2010 operations, refer to the Annual 2010 Report, page 34; for the European Bank for Reconstruction and Development the figure refers to its 2009 operations, refer to the Annual 2009 Report, page 16-20; for the Inter American Development Bank the figure refers to its 2010 operations, refer to the Annual 2010 Report, page 15-25.

These results lead us to suggest that high development performance cannot be achieved unless higher investment on project screening and elaboration is pursued. In our opinion, a substantial message emerging from this paper states clearly that projects need to be solid and well-designed since their proposal, otherwise they do not succeed in terms of development outcome even in presence of good implementation. Throughout this paper we will often refer to "selectivity" as a key term, with this we intend the selection of projects presenting good design and solid structural features. In the empirical analysis we find a clear separation between: ex-ante weak projects, that show very low probability of success even in presence of good implementation; and ex-ante well-designed projects, which have a good chance of succeeding conditional on good implementation. Therefore in this work our reference to selectivity does not recall that of David Dollar and Victoria Levin (2006), whom defined "selective" in terms of democracy and property rights and rule of law.

This paper is structured as follows: in the second section we present the empirical strategy, in the third the data and in the forth the econometric methods; in the fifth section the results are reported and a number of robustness checks are offered; finally in the sixth section we terminate with some concluding remarks.

2 Empirical Strategy

Let us imagine to be in "the best of all possible worlds", in which each project is independently and rigorously evaluated given perfect information and unbiasedness, so that a certain independent agency is in charge of constructing a number of project variables. For example, one may report whether a project was successful (a binary measure), and another variable "how much" it was successful, rating in a given ordinal scale (a multiple measure). Given this imaginative world, let us also pretend that we dispose of the same variables for the quality of the implementer (borrower) and for the agent in charge of supervision and preparation (bank).

Given these pieces of information, we can try to estimate a simple model, like the following:

$$Y_{it} = \alpha\, Borr_{it} + \gamma\, X_{it} + \eta_i + \eta_t + \varepsilon_{it} \quad (1)$$

in which equation (1) explains project success (Y_{it}, in which i indexes the country and t the time) through a simple linear model, which includes as explanatory variables the quality of borrower implementation ($Borr_{it}$), some controls (X_{it}) and fixed effects (η_i, η_t).

The main objective of this paper lies in estimating the borrower performance impact on project outcome, which means estimating in a causal and unbiased way the α coefficient introduced in (1). While we introduce controls for several variables, we are aware that there may be some project-level unobservables, which may covariate with borrower performance. In order to address this problem, our empirical strategy focuses on the adoption of a two-stage estimation, in which for every project we construct some instruments, which are correlated with borrower performance, without being correlated with project specific unobserved factors.

For this reason, we present another equation which interacts with (1)

$$Borr_{it} = \beta\, Bank_{it} + \psi\, Z_{it} + \delta\, X_{it} + \eta_i + \eta_t + u_{it} \quad (2)$$

in which we explain borrower performance ($Borr_{it}$) through bank performance ($Bank_{it}$), a set of other instruments (Z_{it}) and all of the controls and fixed effects introduced in (1).

We are implicitly assuming a transmission channel which links bank to borrower performance, without this variable affecting directly project outcome. Intuitively the channel may be depicted as follows:

Bank Performance → Borrower Performance → Project Outcome

In fact bank performance has an impact on project outcome only if it is able to foster the performance of the borrower, for this reason the two-stage methodology may seem the best way to estimate the previous equation.

Why is the variable ($Bank_{it}$) appropriate for this estimation? The World Bank is a lending institution, which also provides advisory and technical assistance services. It is often referred as a "knowledge bank" because offers expertise in most fields involved in the broad "development business": from telecommunication engineers to water specialists, from accountants to lawyers, from health to energy experts. However the World Bank does not directly implement projects, but provides funding to external entities (national or local governments, private or public companies…) which materially work at their realization. For this reason the Bank agrees on the initial project conditions and then technically follows their implementation, providing supervision and advisory. However, because the Bank does not physically execute the work, its performance is likely not to suffer from project-specific characteristics and thus could be a good instrument for our analysis.

However, this variable may not be sufficient in disentangling the unobservables correlated with the error term and thus for each project we construct four additional instruments:

1. the average borrower and bank performance for the same sector in all countries (except i), in this way we rule out country-specific effects interacting with project outcome (the variables are referred as borr1 and bank1);
2. the average borrower performance for all sectors (except the one in analysis) in all countries (except i), as this cuts off sector-specific effects interacting with project outcome (borr2 and bank2).

This procedure is useful for our purposes and is common in the literature, for example Guasch, Laffont and Straub (2006) successfully adopt it.

3 Data and Literature Review

In the previous section we had to introduce Leibniz's fictitious "the best of all possible worlds" scenario to imagine that projects are evaluated independently and variables are available. Actually this procedure is rather standard in several MDBs and at the same World Bank. This institution presents a whole department dedicated to evaluating projects and other Bank products: the Independent Evaluation Group (IEG, formerly known as OED, Operations Evaluation Department). Its evaluation goals intend to provide an objective assessment of the results of the World Bank's work, through independent and rigorous analysis[2]. Though these data may be criticized and discussed, we feel that they embody important pieces of information, which may lead to useful policy recommendations.

In this part, we would also like to stress that a few papers have already used IEG data to research on policy-relevant issues, though with different research questions and methodologies[3]. For example Isham, Kaufmann and Pritchett (1997) and Isham and Kaufmann (1999) link a specific indicator of project performance[4] with country-level variables, respectively civil liberties and sound macroeconomic policies. Other papers connect project performance to country-wide variables, including institutional quality (Dollar and Levine 2005), conflict (Chauvet et al 2010) and volatility (Guillamont and Laajaj 2006). Another line of research, attempts to explain project-level data with project-level observations: Deininger, Squire, and Basu (1998) analyze the synergies between project success and knowledge/advisory products produced by the World Bank[5]; Pohl and Mihaljek (1998) analyze the degree of project uncertainty between calculated ex-ante and realized ex-post project performances (ERR); while both Dollar and Svensson (2000), Kilby (2000) and Chauvet et al (2006) disentangle the impact of project preparation and

[2] In Annex B we explain the reasons why we consider these ratings to be realistic, serious, reliable and credible.

3 We would like to thank Aart Kraay for the exhausting references provided.

[4] They use the Economic Rate of Return (ERR), which is a measure of the estimated (ex-ante ERR) and realized (ex-post ERR) of economic impact of the project. It is generally defined as the discount rate at which a stream of costs and benefits has a net present value of zero.

5 The expression knowledge products or advisory products refers to the Economic and Sector Works (ESW) and Non Lending Technical Assistance (NLTA): these are "background papers" which provide the World Bank project staff with analytical and advisory documentation, which may be useful in drafting the project, its aims and its implementation.

supervision time on a variety of project-level performance indicators. In a very recent paper, by analyzing project-level variables, Denizer, Kaufman and Kraay (2011) find out that micro variables such as project early-warning indicators, supervision and evaluation lags are all significantly correlated with project outcomes. Another recent work by Fardoust and Flanagan (2011) analyzes the relation between project outcome and the impact of the analytical works developed by the World Bank.

Among the documents and data prepared by this department there is the "Historical Record of the IEG Project Ratings", a database of ratings that IEG has been yearly collecting since 1977. This contains around 9500 projects for several evaluation variables[6], among which there are the ones[7] we are interested in:

- IEG Outcome (from now on Outcome) - it is defined as the extent to which the operation's major relevant objectives were achieved efficiently.
- IEG Overall Borrower Performance (Borrower) - it measures the degree to which the borrower (including the government and implementing agency or agencies) ensured quality of preparation and implementation, and complied with covenants and agreements, toward the achievement of development outcomes.
- IEG Overall Bank Performance (Bank) – it embodies the level to which services provided by the Bank ensured quality at entry of the operation and supported effective implementation through appropriate supervision.

All of these are rated with a multidimensional index, which takes into account different characteristics[8]. Each rating results in a six scale measure (which goes from highly unsuccessful to highly successful)[9] which may be interpreted for econometric purposes as a binary measure

[6] Refer to Annex B for a detailed explanation.

[7] A more rigorous description of the IEG variables can be found at the following links http://web.worldbank.org/external/default/main?theSitePK=1324361&pagePK=64253958&contentMDK=20789730&menuPK=5039271&piPK=64252979
http://web.worldbank.org/external/default/main?theSitePK=1324361&pagePK=64253958&contentMDK=20789685&menuPK=5039271&piPK=64252979

[8] For example, the outcome variable is defined through the following sub-ratings: relevance, efficiency and efficacy, which are themselves divided in other ratings; the same is true for Bank performance (composed by quality at entry and supervision) and Borrower performance (quality of preparation, implementation, covenants compliance and development outcomes).

[9] For the sake of clarity these measures are rated as follows: Highly Satisfactory, Satisfactory, Moderately Satisfactory, Moderately Unsatisfactory, Unsatisfactory, Highly Unsatisfactory.

(with 0 embodying the "unsatisfactory variables" and 1 being the "satisfactory variables")[10] and as a 1 to 6 scale (with 1 being Highly Unsatisfactory and 6 Highly Satisfactory) for the multiple ratings.

Given our definition of infrastructure project[11], 1912 classify for our analysis. They are developed between 1979 to 2008 and represent 136 countries; in Table 2 it is possible to find some summary statistics for the whole dataset, while in Annex A it is possible to find a table reporting the countries in analysis and the number of projects per country. However, among this dataset, only 1710 projects classify for our analysis and present all of the needed ratings.

In this work we use both the binary and the 1-6 ratings. The benefits of the former are several: less sophisticated econometric methodologies, more intuitive results and a simpler analytical framework. However when the interactions between the performance of the project, borrower and supervising agent (bank) are to be studied, we need to disentangle the marginal improvements that performance changes can produce in boosting the project effectiveness; for this reason we employ the 1-6 ratings in order to be as detailed as possible.

Some authors argued that the project, borrower and bank 1-6 ratings may be highly correlated and that it is uncertain to what extent each rating reports independent information. After having read a congruous amount of reports and discussed with evaluators, we feel that this may not be the case as each indicator is separately developed, argued and strongly motivated by facts, interviews and comparisons[12]. For what concerns the correlations, it is indisputable that because we are dealing with ratings, there will be a significant correlation and Table 4 reports the

[10] The satisfactory variables take value 1 and are Highly Satisfactory, Satisfactory, Moderately Satisfactory; while the unsatisfactory variables take value 0 and are Moderately Unsatisfactory, Unsatisfactory, Highly Unsatisfactory.

[11] We define as infrastructure those projects which are classified in the following sectors: Agriculture and Rural Development, Energy and Mining, Environment, Global Information/Communications Technology, Transport, Urban Development and Water. We recognize that also projects classified in excluded sectors may include infrastructure investment (excluded sectors are Economic Policy, Education, Financial and Private Sector Development, Financial Sector, Health Nutrition and Population, Poverty Reduction, Public Sector Governance, Social Development and Social Protection), however we needed define infrastructure projects and thus decided to take a sector board classification. It may be argued that Agriculture and Rural Development may not be considered a fully "infrastructure sector". We recognize that and in our analysis we have also repeated the same exercise excluding this sector, without any significant recorded change. However because in this specific sector, for example, several projects deal with rural electrification, rural transport, irrigation, drains, embankments and other infrastructural improvements within rural development, we strongly believe that this sector should be considered within a broader definition of infrastructure.

[12] Refer to Annex B for further explanations.

correlation coefficients between the three variables[13]. However, the relation between bank, borrower and project performance is the core behind this investigation and as Figure 2 shows through a three-dimensional scatterplot, there is a remarkable dispersion and variability of projects ratings between these variables, which encourages this research.

In addition to these, we also introduce another available project specific control: the mean-normalized size of the project in terms of million dollar value[14]. In addition to this specific control, other country-wide variables are included as well in (X_{it}): GDP per capita (constant 2005 international dollars), from the World Development Indicators (World Bank 2010); the growth rate of income per capita (World Bank 2010) between the years of the project; a measure of volatility, as the standard deviation of GDP per capita (World Bank 2010) during the years of the project, in order to catch the relation between a volatile economy and project performance (Guillamont and Laajaj 2006); an infrastructure index, based on Calderon, Moral-Benito and Serven (2011)[15]; drawing from the previous literature, democratic accountability (reported from now on as accountability) and government stability (reported as stability); some variables that account for the quality of national administrative institutions, bureaucracy and corruption; last but not least that ethnic tensions (reported as ethnic) and external conflict (reported as ext. conflict) as a solid literature (Easterly and Levine, 1997) proofed their detrimental impact on growth.

These last six institutional variables[16] come from the International Country Risk Guide (Political Risk Services 2010), all of them are measures of risk and are constructed through an inverse index: therefore a low measure of "corruption" means a high risk of corruption and viceversa; therefore zero is the highest corruption risk, while six is the lowest.

[13] To clarify this refer to Figure 1 and Annex B as they provide a project-per-project sense of how the ratings are correlated in terms of bank and borrower performance through a rich scatterplot.

[14] We transform the project sizes in constant 2005 international dollars in order to use inter-temporally comparable figures. In addition to this we normalize the project size according to a cross-country inter-temporal mean (104.68 million 2005 PPP dollars). In this way we measure the elasticity of project outcome to relative size, rather than dollar value and it should allow an easier interpretation

[15] From Calderon, Moral-Benito and Serven (2011): "We use a principal component procedure to build a synthetic index summarizing different dimensions of infrastructure. We focus on three key infrastructure sectors: telecommunications, power and road transport. This choice is consistent with previous literature on the output impact of infrastructure, which has typically focused on one of these individual sectors, most often telecommunications. The synthetic infrastructure index is the first principal component of three variables measuring the availability of infrastructure services in these three sectors".

[16] All of their definitions are available in Annex C.

4 Econometric Methods

We are interested in studying the variable "project outcome", which is presented as a binary measure (0 as unsuccessful and 1 as successful). As a robustness check we also study the 1-6 outcome variable (from 1 highly unsuccessful to 6 highly successful), because the ordinal construction of this variable may provide some interesting insights. On the contrary, borrower and bank performance are only analyzed as binary variables. There are several reasons behind this decision (measurement error, sample heterogeneity…), however the most important is that, given the introduced methodological limits, this would not bring any additional information. In fact, it would be meaningless to assess the impact of a rise in the borrower performance coefficient from 4 (moderately successful) to 5 (successful), because this would probably involve a different design of the project and its conditions. For this reason, we only treat borrower and bank performance as binary variables. According to the adopted measure (binary versus multiple), there are differences in the econometric issues encountered and also in terms of coefficient interpretation[17].

For this reason, as widely done in the literature (ie. Denizer, Kaufman and Kraay 2011), a "pragmatic approach" is adopted in estimating this relation. We start from the simplest econometric method using binary variables and then we enrich these estimations by progressively adding one degree of complexity at a time.

Section 5 presents the results of this work, which are developed in accordance with the following scheme:

1. First, the simplest estimation of equation (1) is implemented by adopting the linear probability model and binary variables for both project outcome and borrower performance. Though these estimates present a number of problems, they provide a

[17] For example, the OLS model recognizes the existence of ordered data, but violates some of its basic assumption (homoschedasticity, linearity, normality). On the contrary, multinomial models like logit and probit would not recognize the dependent variable ordinal context. While so called ordered multinomial models, such as ordered probit or logit models controlling for ordinality and linearity, but present much less intuitive results and their interpretation creates non negligible problems. First of all, the elasticity estimates are neither to be considered typical elasticities ($^{dy}/_{dx}$) nor probabilities ($prob(^{dy}/_{dx})$), but marginal probability effects, mainly informative of sign and significance. This means that the computation of the elasticity differs for each independent variable for every value of the dependent variable (i.e. the elasticity of the borrower performance for project outcome is different if the project outcome is 1, 2 or 6). This may be a positive feature, contributing to disentangle complexity, rather than a problem.

"rough sketch" of the probability that a change in the status of the borrower (from unsuccessful to successful) affect the status of the project outcome (from 0 to 1).

2. Keeping the same method, we estimate equation (1) by using the multiple measure for project outcome and a binary for borrower performance. This coefficient embodies the probability that a change in the status of the borrower (from 0 to 1) produces a marginal change in project outcome (from 1 to 6).
3. Given the previous results, in section 5.2 a number of robustness checks are introduced. The first simply repeats the estimation of the previous two steps, by imposing a different distribution of errors (logistic through logit and normal through probit); given these methodologies, both the coefficients and the marginal effects are reported (which in this case are not equal, differently the linear probability model).
4. Through the second check we introduce multinomial ordered models, which allow both for different error distributions (logistic and normal) and to account for the ordinality of the project variable. In this estimation the marginal effects are different for every rating and produce some interesting, yet controversial, results.

For each step, we exclude and introduce country and year fixed effects (η_j, η_t), without significant changes. In these four estimations, in addition to the estimation of (1), we always report the two-stage estimations, which includes (2). In order to clarify these steps and provide an intuitive understanding of the progressive econometric sophistication of this paper, Table 3 provides a quick sketch of this section.

Table 3: Econometric Methods

Section	Method	Outcome Measure	Borrower Measure	Method	Table
5.1	Linear Probability Model	Binary	Binary	OLS/FE 2SLS	Table 5 Table 6
5.1	Linear Probability Model	Multiple	Binary	OLS/FE 2SLS	Table 7 Table 8
5.2.1	Logit and Probit	Binary	Binary	OLS/FE 2SLS	Table 9 Table 10
5.2.1	Ordered Logit and Probit	Multiple	Binary	OLS/FE 2SLS	Table 11 Table 12

Note: the column "Section" refers to the part of next chapter which contains the estimations through the method reported in the column "method". The "Outcome Measure" and "Borrower Measure" columns explain what type of rating is used, either binary (0 unsuccessful or 1 successful) or multiple (from 1 very unsuccessful to 6 very successful), for more references on these measures refer to section 3. OLS means Ordinary Least Squares, FE fixed effects method, 2SLS two stage least squares.

Anticipating on the results we notice that across the estimations there are minor differences and the results of 5.1 seem to be pretty robust. The last estimation is more complex, especially for what regards its interpretation, for this reason it is more deeply developed in section 5.2.2.

5 Results

5.1 The Simplest Estimation

In this section we report the linear probability estimations of equation (1) and the two equations (1) and (2), using binary measures both for outcome and borrower (Table 5 and 6) and a multiple measure for outcome and binary for borrower and bank performance (Table 7 and 8). An intuitive feature of these estimations is given by their coefficients, which are directly interpretable as marginal effects (or probabilities). In general the model seems to explain a significant portion of variability as the adjusted R^2 swings around 0.5; while the two stage approach seems to be useful as the instruments are always very significant and, though we can never validate them, the overidentification tests (Sargan and Basman) do not reject the null hypothesis of their validity, both in Table 6 and in Table 8.

By analyzing tables 5 and 6, we notice that borrower performance is the only significant variable across all estimations (including country fixed effects in FE(1) and both country and year fixed

effects in FE(2)). Growth is significant only when fixed effects are excluded and the institutional variables (bureaucracy and accountability) become significant only when fixed effects are included.

As widely expected when borrowers are successful (rating equals 1) there is a very high probability, almost 70 percent, that also projects meet their development objective and are rated as successful. However, once in Table 6 instruments are included in the estimation, this probability rises by 50 percent to almost to one. For this reason a negative bias seems to affect the estimation of α in Table 5 and the five instruments introduced uncover such bias. This may be explained by strategic behaviors implemented by countries in assigning an implementing team to a project: countries do not randomly match borrowers (and their quality) to projects; they are aware of unobservable factors and according to those each team is matched to a certain set of project characteristics. The results of our estimations, in table 5 and 6 but also in the rest of this work, suggests that countries tend to assign "good implementers" to difficult projects, as a consequence this may undermine the rating of borrower performance and through the two-step approach we uncover this bias.

Tables 7 and 8 confirm these results once multiple ratings are introduced for the outcome variable. Given the rating of a project, a change in the status of a borrower from unsuccessful (0) to successful (1) produces an increase in the rating of roughly 2 points, as Table 7 reports. Also in this case, when instruments are introduced α grows by roughly 50 percent a reaches almost 3.

It is important to underline that in these, like in all other estimations, the impact of bank performance seems to be extremely important in boosting borrower performance. In fact both in table 6 and 8, a change in the status of bank performance increases by 70 percent the probability of a borrower to become successful. Though exploring this link requires a more careful analysis, this fact may represent a stimuli for further research.

5.2 Robustness Checks

5.2.1 Introducing Alternative Distributions

Given our "progressive sophistication approach", in Tables 9 and 10 we replicate the previous estimations by keeping binary measures but assuming a different distribution of errors; therefore

logit and probit regressions are performed, introducing respectively a logistic and normal distribution of error terms.

The model presents again a relatively high R^2, again around 0.5. The instruments are always very significant and not rejected by the overidentification tests. For what regards the estimation results, they confirm exactly the previous section in terms of significance. The same marginal effects are substantially unchanged. In Table 10, we notice that the marginal effects of the borrower variable, in column Logit(1) and Probit(1), are equal to the coefficients reported in Table 5 (~0.68). Also the introduction of the two stage procedure helps correcting for the negative bias and the α coefficient confirms the results obtained in Table 6 rising from 0.68 (columns Logit(1) and Probit(1)) to 0.89 (in column Probit 2SLS (1)). For this reason, once again, if a borrower becomes successful there is a 90 percent probability that the project also turns successful.

5.2.2 Multinomial Ordered Models

With this estimation in addition to alternative distributions, we introduce multiple ratings for project outcome and thus adopt some multinomial ordered models (ordered logit/probit and 2SLS ordered probit). Table 11 reports the estimation results, while Table 12 shows the marginal effects, which in this case are different for each outcome rating. Before proceeding with the analysis it is important to notice that the results mostly confirm previous estimations: R^2 is lower for ordered logit/probit (~0.21- ~0.25) but higher for the two-stage estimation (~0.50).

The marginal effects in Table 12 report a very interesting and powerful result: when project performance is not successful (rating smaller or equal to 4) the impact of a change in borrower performance status does not have a positive impact on project performance. In fact, when the project outcome variable lies between 1 (very unsuccessful) and 3 (moderately unsuccessful) a change in borrower status has a negligible or even negative impact on project performance. On the contrary, when projects are moderately successful (4) or successful (5), a change in borrower performance can increase the probability of a marginal improvement in project outcome between 3 and 9 percent, given a rating of 4, or between 57 and 63 percent, given a rating of 5.

These results lead to state that successful borrower performance can have an impact on the project, if and only projects are structurally solid and constructed considering how realistic its

success can be. In fact, unsuccessful projects (with a rating between 1 and 3) are never going to turn successful, even if the borrower performance fulfils the requirements and excels. On the other hand, successful projects benefit significantly from improvements in borrower performance and thus can reach the highest possible development outcome, conditional on excellent project implementation.

It may be argued that when projects are moderately successful (rating of 4), the change in the status of a borrower does produce small or even insignificant effects on project outcome. This result may be significantly biased by the fact that a number of these projects (with a rating of 4), should have lower rating (3 for example), but are inflated and thus affect the significance of our coefficients.

6 Conclusion

This work attempts to study the determinants of World Bank project success and disentangle the causality behind the relation between the outcome of projects and the performance of the implementer (borrower).

These results suggest unambiguously that the selection of solid projects, that we call selectivity, is the fundamental tool for successful infrastructure investment: badly designed projects are never going to deliver successful development outcomes, even if implementers comply with quality standards and execute effectively their duty (as Table 12 shows). At the same time, even when "good" projects are approved, the World Bank needs to pick up excellent implementers, because its knowledge and financial support contribute to borrower performance, and thus project outcome, only in this case.

The fact that macroeconomic and institutional variables do not significantly explain project outcome suggests that broad institutional reforms and good macroeconomic management may be useful for improving the business climate, countries' competitiveness and sound budget positions; however in terms of aid and project effectiveness, selectivity is the key and only well-designed and solid projects combined with high-performing implementers are successful.

The policy recommendations are very straightforward: the World Bank needs to strengthen its capacity building programs and place more and more attention on project governance. With this

term we mean the financing of training and development of countries' infrastructure implementing authorities (public or private companies, ministries, agencies…) and finance well-designed and solid projects, conditional on the fact that these authorities achieved high effectiveness.

The first problem with this recommendation is that an "equity versus effectiveness" dilemma may emerge, because "effective borrowers" may be found in those countries who need the Bank support the least. However this needs to be soon demystified: only successful projects deliver significant development results, which help fighting poverty and promoting development. For this reason such dilemma does not really hold and, in our opinion, the World Bank and other development institutions should place attention on effectiveness and results. A serious problem which may prevent this strategy from being implemented resides on the incentive structure that Multilateral Development Banks, and the World Bank itself, present: providing funding quickly in response to countries' needs may be much more valued than a careful job of project screening and elaboration, researching on countries' true priorities and sectors that unbundle development. This may contrast with the "fast lending" desire that the Bank may use as a poor quantitative productivity criterion; this pressure on lending speed may be exactly what some corrupted officials are keen on, because interested in the lucrative earnings behind loan, rather than its development outcome. In our opinion internal reforms which aim to the alignment of incentives toward promoting selectivity and developing more project-specific a-priori analysis are key to develop successful projects.

For what regards World Bank performance, it must be stated that we dispose of little evidence regarding its determinants and in this research we only preliminarily studied it. From the literature, it seems that some reforms had an impact as they promoted a more adequate team composition and a balanced allocation of time between project preparation and supervision (Denizer, Kaufmann and Kraay, 2011). There is also another interesting channel, which links the project ratings with the quality of the knowledge products[18] involved in their project preparation (Fardoust and Flanagan, 2011): this may be central because as the World Bank becomes a knowledge bank, the "knowledge intensity" of its projects is due to rise and thus eventually promote more performance gains. For all of these reasons, we feel that more data collection

[18] Refer to the definition provided in footnote 5.

needs to be implemented in order to understand what drives the World Bank performance, how Bank performance affects borrowers and project, and finally which reforms produced a real enhancing effect.

Though external validity warnings are to be made, we believe that the World Bank experience may be important for other multilateral development banks, international development institutions and, even, national governments. Selectivity is the key and every effort to promote development through infrastructure investment need to focus on picking up well-designed and solid projects: this may require a huge investment by financing institutions in project elaboration and selection, but this seems to be the most important micro step toward the realization of a "world free of poverty".

References

Agenor, Pierre R. and Moreno-Dodson, Blanca. (2006). "Public Infrastructure and Growth: New Channels and Policy Implications", World Bank Policy Research Working Paper No. 4064.

Calderon, Cesar, Moral-Benito, Enrique and Serven, Luis. (2011). "Is infrastructure capital productive? A dynamic heterogeneous approach", Banco de España Working Papers 1103.

Chauvet, Lisa, Paul Collier, and Margeurite Duponchel (2010). "What Explains Aid Project Success in Post-Conflict Situations?". World Bank Policy Research Working Paper No. 5418.

Chauvet, Lisa, Paul Collier, and Andreas Fuster (2006). "Supervision and Project Performance: A Principal-Agent Approach". Manuscript, DIAL.

Deininger, Klaus, Lyn Squire, and Swati Basu (1998). "Does Economic Analysis Improve the Quality of Foreign Assistance?". World Bank Economic Review. 12(3):385-418.

Denizer, Cevit, Kaufmann and Kraay, Aart. (2011). "Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance", Policy Research Working Paper No. 5646.

Dollar, David and Jakob Svensson (2000). "What Explains the Success and Failure of Structural Adjustment Programs?". The Economic Journal. 110():894-917.

Dollar, David and Victoria Levin (2005). "Sowing and Reaping: Institutional Quality and Project Outcomes in Developing Countries". World Bank Policy Research Working Paper No. 3524.

Dollar, David and Victoria Levin (2006). "The Increasing Selectivity of Foreign Aid: 1984-2003", World Development, Vol. 34, No. 12, pp. 2034–2046.

Estache, Antonio and Marianne Fay. (1997). " Ensuring regional growth convergence in Argentina and Brazil: How can governments help?", mimeo, The World Bank.

Easterly, William and Levine, Ross. (1997). "Africa's Growth Tragedy: Policies and Ethnic Divisions", The Quarterly Journal of Economics, Vol. 112, No. 4, Pages 1203-1250.

Fardoust, Shahrokh and Flanagan, Ann Elizabeth. (2011). "The Quality of Economic & Sector Analysis and the World Bank Loan Ratings: Any Statistical Relationship?". Policy Research Working Paper, forthcoming.

Guillaumont, Patrick and Rachid Laajaj (2006). "When Instability Increases the Effectiveness of Aid Projects". World Bank Policy Research Working Paper No. 4034.

Isham, Jonathan and Daniel Kaufmann (1999). "The Forgotten Rationale for Policy Reform: The Productivity of Investment Projects". Quarterly Journal of Economics. 114(1):149-184

Isham, Jonathan, Daniel Kaufmann and Lant Pritchett (1997). "Civil Liberties, Democracy, and the Performance of Government Projects". World Bank Economic Review. 11(2): 219-242.

Kilby, Christopher (2000). "Supervision and Performance: The Case of World Bank Projects". Journal of Development Economics. 62: 233-259.

Pohl, Gerhard and Dubravko Mihaljek (1998). "Project Evaluation and Uncertainty in Practice: A Statistical Analysis of Rate-of-Return Divergences in 1015 World Bank Projects". World Bank Economic Review. 6(2): 255-257.

Political Risk Services. (2010). "International country risk guide". New York: Political Risk Services, 2010.

Röller, Lars-Hendrik and Waverman, Leonard, 2001, "Telecommunications Infrastructure and Economic Development: A Simultaneous Approach", The American Economic Review, Vol. 91, No. 4 (Sep., 2001), pp. 909-923.

Sigurdsson, S. and Schweitzer, E. 1995. "Performance indicators in Bank-financed operations: second edition". Human Development Department, Working Paper, December 1995.

Wapenhans, Willi A. (1992). "Report on the task force on portfolio management", The World Bank.

World Bank (2010). "World Development Indicators", Development Economic Vice Presidency. The World Bank Group, Washington DC.

World Bank (1994). "World Development Report: Infrastructure for Development". The World Bank Group, Washington DC.

Figure 1: rich scatterplot of Bank and Borrower variables

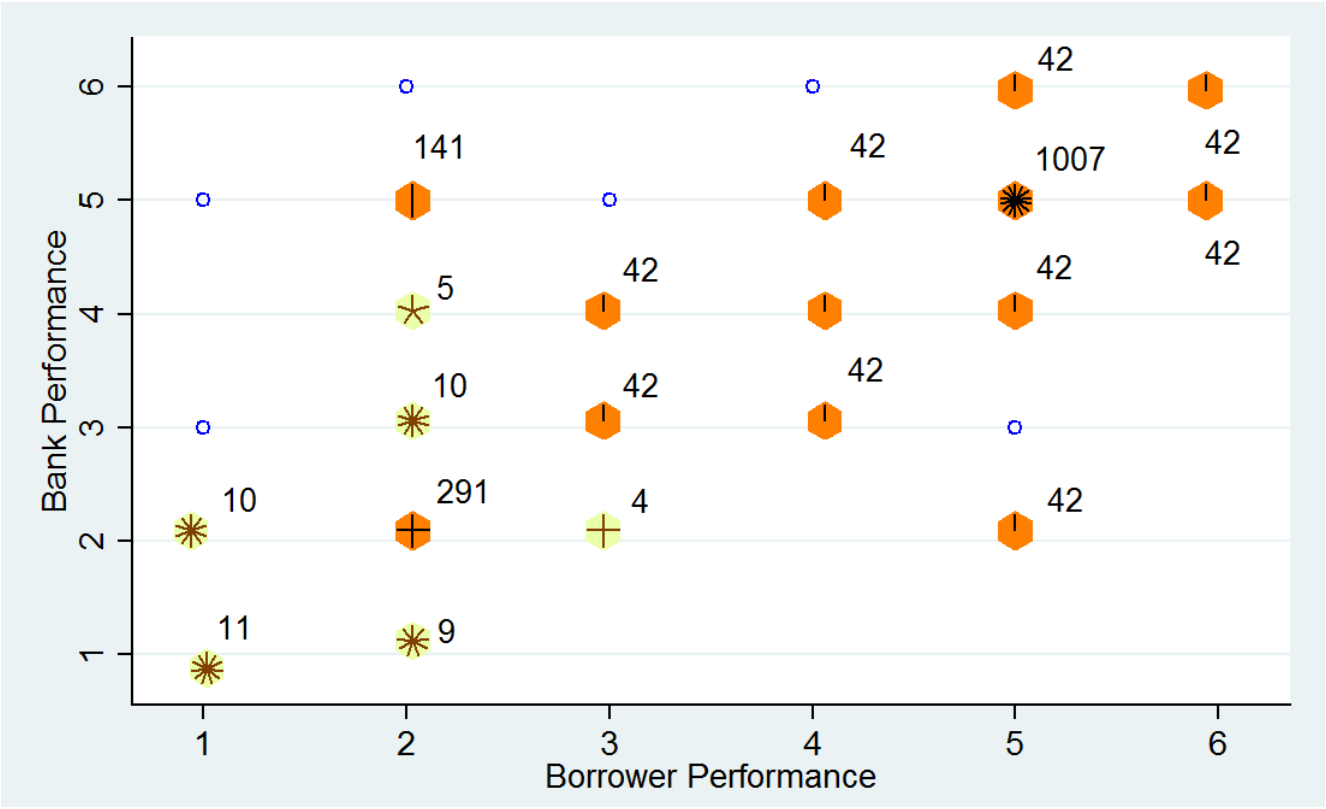

Figure 2: 3d scatterplot of Project Outcome, Bank and Borrower Performance

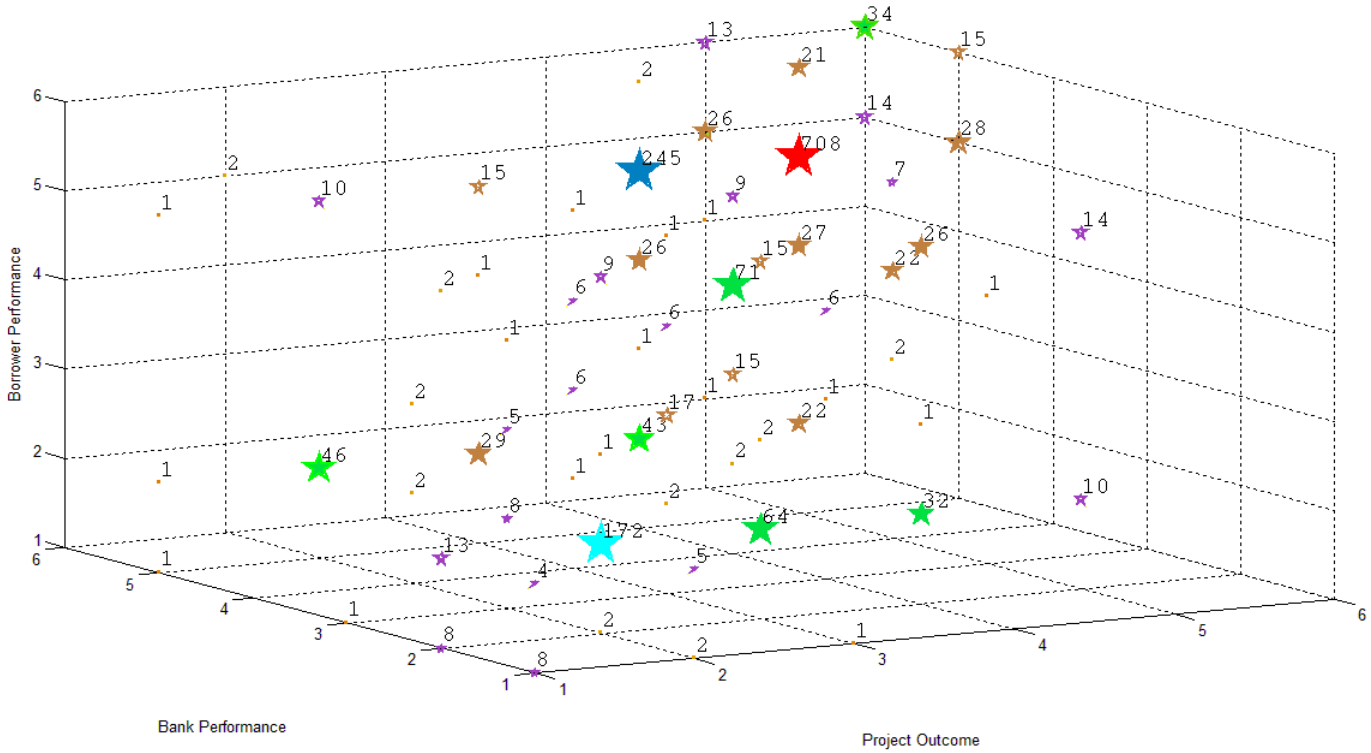

TABLE 2: summary statistics

	Observations	Mean (st. dev)	Min	Max
Outcome Multiple	1912	4.127 (1.207)	1	6
Outcome Binary	1912	0.754 (0.430)	0	1
Borrower Multiple	1912	4.140 (1.354)	1	6
Borrower Binary	1912	0.722 (0.447)	0	1
Bank Multiple	1912	4.296 (1.282)	1	6
Bank Binary	1912	0.769 (0.421)	0	1
Size	1912	1 (1.209) [104.088]*	0	8.988
GDP p.c.	1891	4177.30 (3678.40)	351.03	21211.91
Volatility	1866	602.14 (726.80)	0	4948.75
Growth	1872	0.247 (0.314)	-1	2.353
Infrastructure	1891	0.053 (0.063)	0.001	0.321
Bureaucracy	1742	1.944 (0.730)	0	4
Corruption	1742	2.568 (0.739)	0.5	5
Accountability	1742	3.439 (1.333)	0	6
Stability	1742	8.179 (1.607)	2.5	11.5
Ethnic	1742	3.744 (1.202)	0	6
Ext. conflict	1742	9.942 (1.391)	3.5	12

Note: the asterisk indicates value reported in square brackets, which is the mean size of all projects available in our database, reported in 2005 constant international dollars.

TABLE 4: correlation coefficients for binomial and six scale variables

Binomial variables

	Outcome	Borrower	Bank
Outcome	1		
Borrower	0.731	1	
Bank	0.695	0.658	1

Six scale variables

	Outcome	Borrower	Bank
Outcome	1		
Borrower	0.758	1	
Bank	0.707	0.688	1

TABLE 5: regressions on binary ratings of outcome and borrower

	OLS (1)	FE (1)	FE (2)	
	Binary Outcome	Binary Outcome	Binary Outcome	
Borrower	0.688***	0.687***	0.683***	
	(0.016)	(0.016)	(0.017)	
Size	-0.008	-0.010	-0.009	
	(0.006)	(0.007)	(0.007)	
GDP p.c.	0.000005	-0.000006	-0.000001	
	(0.00004)	(0.00004)	(0.00002)	
Volatility	-0.00002	-0.00001	-0.00001	
	(0.00002)	(0.00002)	(0.00002)	
Growth	0.104***	0.075	0.057	
	(0.040)	(0.062)	(0.067)	
Infrastructure	0.139	-0.014	-0.088	
	(0.212)	(0.046)	(0.483)	
Bureaucracy	-0.006	-0.045**	-0.050**	
	(0.011)	(0.023)	(0.023)	
Corruption	-0.002	-0.017	-0.011	
	(0.010)	(0.017)	(0.018)	
Accountability	-0.010*	0.032**	0.027**	
	(0.005)	(0.012)	(0.012)	
Stability	0.001	-0.003	-0.002	
	(0.005)	(0.006)	(0.012)	
Ethnic	0.005	0.021	0.017	
	(0.006)	(0.016)	(0.017)	
Ext. conflict	0.003	-0.005	-0.007	
	(0.005)	(0.008)	(0.009)	
Country FE		Yes	Yes	
Year FE			Yes	

Other Statistics

Observ.	1710	1710	1710	
Adj. R^2	0.545	0.557	0.556	
Prob(F)	0.00	0.00	0.00	

Note: both the project outcome variable and the borrower performance variable are reported on a 0 (unsuccessful) and 1 (successful) scale. For further references, see the text.

TABLE 6: regressions on binary ratings of outcome and borrower

	2SLS (1)		2SLS (2)		2SLS (3)	
	Outcome	Borrower	Outcome	Borrower	Outcome	Borrower
Borrower	1.009***		1.027***		1.027***	
	(0.027)		(0.029)		(0.029)	
Size	-0.007	-0.009	-0.007	-0.001	-0.005	0.007
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)
GDP p.c.	-0.00001	0.00001	-0.00004	0.00004	-0.00001	0.00001
	(0.0005)	(0.00005)	(0.0002)	(0.00002)	(0.0002)	(0.00002)
Volatility	-0.00001	-0.00001	-0.00003	0.00003	-0.00003	0.00004
	(0.00002)	(0.0002)	(0.00003)	(0.00003)	(0.00003)	(0.00003)
Growth	-0.002	0.003	0.052	0.035	0.032	0.036
	(0.052)	(0.059)	(0.087)	(0.086)	(0.093)	(0.095)
Infrastructure	0.265	0.324	-0.645	0.300	-0.805	0.334
	(0.266)	(0.279)	(0.614)	(0.648)	(0.639)	(0.657)
Bureaucracy	0.007	-0.0004	-0.046*	0.005	-0.040	0.005
	(0.015)	(0.013)	(0.025)	(0.032)	(0.026)	(0.033)
Corruption	-0.010	0.019	-0.020	-0.007	-0.014	-0.004
	(0.014)	(0.015)	(0.024)	(0.025)	(0.025)	(0.026)
Account.	-0.006	-0.010	0.041***	-0.024*	0.039***	-0.026*
	(0.006)	(0.007)	(0.014)	(0.014)	(0.014)	(0.015)
Stability	-0.013*	0.003	-0.009	0.005	-0.016*	0.002
	(0.007)	(0.007)	(0.007)	(0.010)	(0.009)	(0.010)
Ethnic	-0.001	-0.003	0.004	-0.009	0.005	-0.011
	(0.008)	(0.008)	(0.023)	(0.023)	(0.023)	(0.024)
Ext. conflict	-0.005	0.004	-0.016*	0.021**	-0.015	0.018*
	(0.007)	(0.008)	(0.009)	(0.009)	(0.011)	(0.011)
Instruments						
Bank		0.667***(0.02)		0.661***(0.02)		0.661***(0.02)
Bank 1		3.48***(0.61)		2.53***(0.77)		2.58***(0.79)
Bank 2		16.3***(2.48)		12.3***(3.30)		12.1*** (3.36)
Borr 1		-4.49***(0.74)		-3.74*** (1.05)		-3.84*** (1.10)
Borr 2		-21.9***(3.15)		-19.0*** (4.75)		-18.9*** (4.96)
Country FE			Yes	Yes	Yes	Yes
Year FE					Yes	Yes

Other Statistics

Observ.	1710	1710	1710	1710	1710	1710
Adj. R^2	0.436	0.462	0.439	0.470	0.440	0.470
Prob(F)	0.00	0.00	0.00	0.00	0.00	0.00
Sargan Test	0.116		0.741		0.635	
Basman Test	0.118		0.766		0.673	

Note: both the project outcome variable and the borrower performance variable are reported on a 0 (unsuccessful) and 1 (successful) scale. For further references, see the text. The number of observations (1709) is lower than the sample because some instruments do not present a total coverage of the sample.

TABLE 7: regressions on multiple outcome and binary borrower

	OLS (1)	FE (1)	FE (2)	
	Multiple Outcome	Multiple Outcome	Multiple Outcome	
Borrower	1.915***	1.905***	1.896***	
	(0.096)	(0.047)	(0.048)	
Size	0.012	-0.025	-0.025	
	(0.017)	(0.019)	(0.019)	
GDP p.c.	-0.000007	-0.000001	-0.000001	
	(0.00001)	(0.00001)	(0.00001)	
Volatility	0.0001	-0.00005	-0.00005	
	(0.0001)	(0.00002)	(0.00002)	
Growth	0.327*	0.116	0.116	
	(0.113)	(0.179)	(0.179)	
Infrastructure	0.512	-0.000	-0.000	
	(1.173)	(0.475)	(0.475)	
Bureaucracy	0.105	-0.120*	-0.157*	
	(0.077)	(0.065)	(0.068)	
Corruption	0.004	-0.026	-0.026	
	(0.077)	(0.017)	(0.017)	
Accountability	0.034	0.115***	0.108***	
	(0.031)	(0.035)	(0.036)	
Stability	0.027	-0.001	-0.001	
	(0.038)	(0.006)	(0.006)	
Ethnic	0.009	0.015	0.015	
	(0.041)	(0.016)	(0.016)	
Ext. conflict	0.037*	-0.007	-0.007	
	(0.015)	(0.008)	(0.008)	
Country FE		Yes	Yes	
Year FE			Yes	
Other Statistics				
Observ.	1710	1710	1710	
Adj. R^2	0.534	0.550	0.552	
Prob(F)	0.00	0.00	0.00	

Note: both the project outcome variable and the borrower performance variable are reported on a 0 (unsuccessful) and 1 (successful) scale. For further references, see the text.

TABLE 8: regressions on multiple outcome and binary borrower

	2SLS (1)		2SLS (2)		2SLS (3)	
	Outcome	Borrower	Outcome	Borrower	Outcome	Borrower
Borrower	2.750***		2.755***		2.766***	
	(0.077)		(0.082)		(0.083)	
Size	-0.009	0.028	-0.009	0.028	-0.008	0.028
	(0.009)	(0.029)	(0.009)	(0.029)	(0.009)	(0.029)
GDP p.c.	-0.00001	0.00005	-0.00001	0.00005	-0.00001	0.00001
	(0.00002)	(0.00007)	(0.00002)	(0.00007)	(0.0002)	(0.00007)
Volatility	-0.00001	0.00005	-0.00001	0.00005	-0.00003	0.00007
	(0.00003)	(0.0001)	(0.00003)	(0.0001)	(0.00003)	(0.00001)
Growth	0.119	-0.221	0.015	-0.221	0.116	-0.145
	(0.124)	(0.265)	(0.196)	(0.265)	(0.095)	(0.285)
Infrastructure	-0.550	0.863	-0.550	0.863	-0.836	1.113
	(0.630)	(1.942)	(0.630)	(1.942)	(0.654)	(1.964)
Bureaucracy	-0.014	0.004	-0.122**	0.004	-0.116	0.003
	(0.035)	(0.098)	(0.071)	(0.098)	(0.074)	(0.099)
Corruption	-0.020	0.011	-0.020	0.011	-0.028	0.024
	(0.018)	(0.075)	(0.018)	(0.075)	(0.026)	(0.077)
Account.	0.003	-0.001	0.137***	-0.023*	0.148***	-0.026*
	(0.017)	(0.006)	(0.039)	(0.014)	(0.041)	(0.015)
Stability	-0.007	0.002	-0.007	0.002	-0.007	0.015
	(0.009)	(0.031)	(0.009)	(0.031)	(0.010)	(0.031)
Ethnic	0.010	0.012	0.010	0.012	0.004	-0.003
	(0.024)	(0.071)	(0.024)	(0.071)	(0.024)	(0.072)
Ext. conflict	-0.021	0.057*	-0.005	0.021***	-0.004	0.018*
	(0.017)	(0.029)	(0.027)	(0.009)	(0.030)	(0.011)
Instruments						
Bank		0.667***(0.02)		0.661***(0.02)		0.661***(0.02)
Bank1		3.48***(0.60)		2.53***(0.77)		2.58***(0.79)
Bank2		16.3***(2.48)		12.3***(3.30)		12.1***(3.36)
Borr1		-4.49***(0.74)		-3.74***(1.0)		-3.84***(1.0)
Borr2		-21.9***(3.1)		-19.0***(4.78)		-18.9***(4.96)
Country FE			Yes	Yes	Yes	Yes
Year FE					Yes	Yes

Other Statistics

Observ.	1710	1710	1710	1710	1710	1710
Adj. R^2	0.445	0.462	0.461	0.470	0.456	0.470
Prob(F)	0.00	0.00	0.00	0.00	0.00	0.00
Sargan Test	0.182		0.479		0.391	
Basman Test	0.185		0.516		0.418	

Note: both the project outcome variable and the borrower performance variable are reported on a 0 (unsuccessful) and 1 (successful) scale. For further references, see the text. The number of observations (1709) is lower than the sample because some instruments do not present a total coverage of the sample.

TABLE 9: logit, probit and ivprobit regressions on binomial outcome and borrower

	Logit (1)	Logit FE (2)	Probit (1)	Probit FE (2)	Probit 2SLS (1)	
	Outcome	Outcome	Outcome	Outcome	Outcome	Borrower
Borrower	4.079***	4.766***	2.321***	2.653***	4.621***	
	(0.181)	(0.246)	(0.092)	(0.119)	(0.249)	
Size	-0.101	-0.145	-0.051	-0.072	-0.066	0.005
	(0.074)	(0.093)	(0.038)	(0.049)	(0.082)	(0.010)
GDP p.c.	-0.00006	-0.00002	-0.00003	-0.0001	-0.0001	-0.00002
	(0.0005)	(0.0003)	(0.0002)	(0.0001)	(0.0002)	(0.0003)
Volatility	-0.0003	-0.0001	-0.00001	-0.00003	-0.0003	0.00004
	(0.0002)	(0.0004)	(0.0001)	(0.0002)	(0.0003)	(0.00004)
Growth	1.479***	0.771	0.781***	0.464	1.499	0.009
	(0.536)	(1.078)	(0.280)	(0.562)	(0.936)	(0.105)
Infrastructure	1.893	3.231	1.077	1.138	-3.782	0.558
	(2.907)	(8.501)	(1.501)	(4.545)	(7.132)	(0.796)
Bureaucracy	-0.085	-0.618**	-0.034	-0.306*	0.153	0.0004
	(0.135)	(0.309)	(0.071)	(0.159)	(0.269)	(0.035)
Corruption	-0.030	-0.197	-0.014	-0.102	-0.248	0.0008
	(0.127)	(0.255)	(0.066)	(0.135)	(0.226)	(0.028)
Accountability	-0.146**	-0.263	-0.081**	-0.131	0.224*	-0.026*
	(0.072)	(0.172)	(0.038)	(0.092)	(0.116)	(0.016)
Stability	-0.016	-0.067	-0.003	-0.036	-0.086	-0.001
	(0.061)	(0.100)	(0.032)	(0.052)	(0.072)	(0.019)
Ethnic	0.051	0.240	0.026	0.094	0.095	0.009
	(0.082)	(0.228)	(0.043)	(0.120)	(0.225)	(0.026)
Ext. conflict	0.033	-0.122	0.015	-0.063	-0.117	0.019
	(0.062)	(0.117)	(0.033)	(0.063)	(0.077)	(0.011)
Instruments						
Bank						0.665***(0.02)
Bank1						2.61***(0.82)
Bank2						12.6***(3.51)
Borr1						-3.88***(1.10)
Borr2						-19.5***(5.17)
Country FE		Yes		Yes	Yes	Yes
Year FE		Yes		Yes	Yes	Yes

Other Statistics

Observ.	1710	1576^{-}	1710	1576^{-}	1582^{-}	
L.Likelihood	-477.50	-396.97	-477.53	-395.97		
Pseudo R^2	0.489	0.553	0.489	0.554	0.467	
Prob(χ^2)	0.000	0.000	0.000	0.000	0.000	
ALN test^{+}					0.293	

Note: $^{-}$ some observations removed because of collinearity between the country and year effects. $^{+}$ALN stands for the Amemiya-Lee-Newey minimum chi-sq statistic, it is an overidentification test performed on the ivprobit using the two-step estimation method.

TABLE 10: marginal effects of logit, probit and ivprobit regressions on binomial outcome of borrower performance

	Logit (1)	Logit FE (2)	Probit (1)	Probit FE (2)	Probit 2SLS (1)[+]
	Outcome	Outcome	Outcome	Outcome	Outcome
Borrower	0.680***	0.741***	0.682***	0.732***	0.897***
	(0.023)		(0.022)		
Size	-0.010	-0.013	-0.011	-0.014	-0.001
	(0.007)		(0.008)		
GDP p.c.	0.00006	-0.00002	0.00007	-0.00002	-0.00009
	(0.0001)		(0.0001)		
Volatility	-0.00003	0.00001	-0.00003	0.00007	-0.00009
	(0.0003)		(0.0003)		
Growth	0.154*	0.067	0.171***	0.089	0.062
	(0.055)		(0.061)		
Infrastructure	0.197	0.295	0.236	0.226	0.138
	(0.302)		(0.329)		
Bureaucracy	-0.008	-0.056**	-0.007	-0.060*	-0.024
	(0.014)		(0.015)		
Corruption	-0.003	-0.019	-0.003	-0.021	-0.003
	(0.013)		(0.014)		
Accountability	-0.015**	0.023	-0.018**	0.024	0.028
	(0.007)		(0.008)		
Stability	-0.001	-0.006	-0.001	-0.006	-0.010
	(0.007)		(0.007)		
Ethnic	0.005	0.022	0.005	0.019	0.020
	(0.008)		(0.009)		
Ext. conflict	0.003	-0.011	0.003	-0.012	-0.010
	(0.006)		(0.007)		
Country FE		Yes		Yes	Yes
Year FE		Yes		Yes	Yes

Note: [+] the marginal effects computed for the IV probit come from a numerical procedure, which requests that all specified variables and observations be retained in the maximization process. The results from this regression do not significantly change from the 2SLS probit, for which however we cannot compute the marginal effects.

TABLE 11: ordered logit, probit and 2SLS ordered probit regressions on multiple outcome and binary borrower

	Ologit (1)	Ologit FE (2)	Oprobit (1)	Oprobit FE (2)	Oprobit 2SLS (1)	
	Outcome	Outcome	Outcome	Outcome	Outcome	Borrower
Borrower	3.904***	4.121***	2.143***	2.727***	2.919***	
	(0.154)	(0.164)	(0.075)	(0.081)	(0.081)	
Size	-0.094	-0.155	-0.051	-0.072	-0.002	0.005
	(0.076)	(0.095)	(0.038)	(0.049)	(0.027)	(0.010)
GDP p.c.	-0.00006	-0.00002	-0.00003	-0.0001	-0.0001	-0.00002
	(0.00006)	(0.0003)	(0.0002)	(0.0001)	(0.0002)	(0.0003)
Volatility	-0.0003	-0.0001	-0.00001	-0.00003	-0.0003	0.00004
	(0.0003)	(0.0004)	(0.0001)	(0.0002)	(0.0003)	(0.00004)
Growth	0.751***	0.771	0.417***	0.464	1.499	0.009
	(0.278)	(1.078)	(0.155)	(0.562)	(0.936)	(0.105)
Infrastructure	1.452	3.231	1.077	1.138	-3.782	0.558
	(2.907)	(8.501)	(1.501)	(4.545)	(7.132)	(0.796)
Bureaucracy	-0.085	-0.362**	-0.034	-0.229**	-0.152*	0.0004
	(0.135)	(0.170)	(0.071)	(0.095)	(0.086)	(0.035)
Corruption	0.138*	-0.197	0.070	-0.102	-0.248	0.0008
	(0.072)	(0.255)	(0.040)	(0.135)	(0.226)	(0.028)
Accountability	0.021	0.235**	0.004	0.152***	0.152***	-0.035
	(0.038)	(0.092)	(0.038)	(0.051)	(0.048)	(0.046)
Stability	-0.016	-0.067	-0.003	-0.036	-0.110*	-0.001
	(0.061)	(0.100)	(0.032)	(0.052)	(0.065)	(0.019)
Ethnic	0.051	0.240	0.026	0.094	0.095	0.009
	(0.082)	(0.228)	(0.043)	(0.120)	(0.225)	(0.026)
Ext. conflict	0.095	0.125	0.047**	-0.063	-0.117	0.012
	(0.037)	(0.065)	(0.021)	(0.063)	(0.077)	(0.032)
Instruments						
Bank						2.18*** (0.09)
Bank1						16.8*** (3.0)
Bank2						73.4*** (12.6)
Borr1						-21.8*** (3.84)
Borr2						-102.8*** (16.4)
Country FE		Yes		Yes	Yes	
Year FE		Yes		Yes	Yes	Yes

Other Statistics

Observ.	1710	1710	1710	1710	1710	
L.Likelihood	-1875.52	-1780.69	-1884.79	-1786.43	-2304.54	
Pseudo R^2	0.216	0.255	0.212	0.253	0.577[#]	
Prob(χ^2)	0.000	0.000	0.000	0.000	0.000	

Note: [−] some observations removed because of collinearity between the country and year effects. [+] In this estimation the 2SLS ordered probit estimaton contains only year fixed effects, because the introduction of country dummies prevents matrix convergence to take place. [#] this R^2 is not a pseudo, like in ordered logit and probit, but unadjusted R^2.

TABLE 12: marginal effects of ordered logit, probit and 2SLS ordered probit regressions on multiple outcome of binary borrower performance

	Ologit (1)	Ologit FE (2)	Oprobit (1)	Oprobit FE (2)	Oprobit 2SLS (1)	
	Borrower	Borrower	Borrower	Borrower	Borrower	
Outcome=1	-0.063***	-0.049***	-0.070***	-0.053***	-0.156***	
	(0.010)		(0.015)			
Outcome =2	-0.417***	-0.423***	-0.411***	-0.423***	-0.528***	
	(0.022)		(0.021)			
Outcome =3	-0.194***	-0.214***	-0.166***	-0.187***	-0.141***	
	(0.016)		(0.014)			
Outcome =4	0.033*	0.027*	0.012	0.011	0.093***	
	(0.019)		(0.017)			
Outcome =5	0.581***	0.608***	0.572***	0.602***	0.638***	
	(0.014)		(0.014)			
Country FE		Yes		Yes	Yes	
Year FE		Yes		Yes	Yes	

Note: [+] the marginal effects computed for the IV probit come from a numerical procedure, which requests that all specified variables and observations be retained in the maximization process. The results from this regression do not significantly change from the 2SLS probit, for which however we cannot compute the marginal effects.

Annex A: Countries and projects involved

Country	Number of Projects	Country	Number of Projects
Afghanistan	6	Kyrgyz Republic	15
Albania	23	Lao PDR	14
Algeria	26	Latvia	8
Angola	4	Lebanon	9
Argentina	21	Lesotho	5
Armenia	11	Lithuania	10
Azerbaijan	10	Macedonia, FYR	6
Bahamas, The	1	Madagascar	18
Bangladesh	35	Malawi	19
Barbados	1	Malaysia	6
Belarus	3	Mali	17
Belize	4	Mauritania	12
Benin	13	Mauritius	6
Bhutan	4	Mexico	40
Bolivia	23	Moldova	8
Bosnia and Herzegovina	18	Mongolia	5
Botswana	1	Montenegro	1
Brazil	97	Morocco	38
Bulgaria	13	Mozambique	23
Burkina Faso	14	Nepal	22
Burundi	9	Nicaragua	11
Cambodia	8	Niger	10
Cameroon	11	Nigeria	24
Cape Verde	2	Pakistan	44
Central African Republic	5	Panama	3
Chad	9	Papua New Guinea	10
Chile	12	Paraguay	6
China	151	Peru	12
Colombia	21	Philippines	39
Comoros	2	Poland	33
Congo, Rep.	3	Romania	22
Costa Rica	5	Russian Federation	21
Cote d'Ivoire	13	Rwanda	8
Croatia	11	Samoa	4
Cyprus	4	Sao Tome and Principe	1
Czech Republic	5	Senegal	18
Djibouti	3	Serbia	3
Dominica	2	Seychelles	2

Dominican Republic	7	Sierra Leone	6
Ecuador	17	Slovak Republic	2
Egypt, Arab Rep.	15	Slovenia	4
El Salvador	5	South Africa	2
Equatorial Guinea	1	Sri Lanka	23
Eritrea	1	St. Lucia	4
Estonia	5	St. Vincent and G.	3
Ethiopia	17	Sudan	1
Fiji	3	Swaziland	1
Gabon	3	Tajikistan	6
Gambia, The	5	Tanzania	25
Georgia	15	Thailand	21
Ghana	37	Timor-Leste	6
Grenada	2	Togo	7
Guatemala	5	Tonga	1
Guinea	14	Trinidad and Tobago	2
Guinea-Bissau	4	Tunisia	21
Guyana	4	Turkey	30
Haiti	5	Turkmenistan	2
Honduras	11	Uganda	27
Hungary	15	Ukraine	13
India	107	Uruguay	14
Indonesia	69	Uzbekistan	5
Iran, Islamic Rep.	8	Vanuatu	1
Jamaica	8	Venezuela, RB	9
Jordan	13	Vietnam	21
Kazakhstan	11	West Bank and Gaza	12
Kenya	27	Yemen, Rep.	25
Korea, Rep.	7	Zambia	14
Kosovo	6	Zimbabwe	7

Annex B: How is the IEG database built? Why are these ratings useful?

The evaluations reported in this database are constructed through two methods: ES/EVM and PPAR. The ES/EVM, evaluation summary or memoranda, are project evaluations prepared and recorded by the responsible Bank department (not by IEG) at the end of the project (reported in the ICR – Implementation Completion Report). These data are then reviewed by the ICR Review board (this procedure was established in 1995), which is an intermediate World Bank-IEG committee which seeks to independently verify the findings of the ICR. The problems behind EVM evaluations are linked with the self-assessment produced by projects responsible, which may create an upward bias in the data. On the contrary, the Project Performance Assessment Report (PPAR) offers a much stronger picture of the evaluation. These are developed by IEG, which independently and randomly samples 25% of World Bank projects per year, sends its experts on field to evaluate the results of these projects and create independent evaluations

Though these ratings may not be perfect, we feel that they contain some important pieces of information, which may provide useful policy recommendations. One could ask whether these evaluation ratings are realistic? We feel they are pretty much. The department within the World Bank in charge of their collection, IEG, spends substantial resources in reviews and quality controls, employs very experienced evaluators, has very strict confidentiality and anti-conflict of interest norms. In addition to this, the institutional architecture of the World Bank structures IEG as a special and independent department, which reports only to the Board of Directors (the decision-making entity of the World Bank) and not to single departments/vice presidencies, it does not share personnel, resources or initiatives with other parts of the Bank; surprisingly enough IEG does not even share the location with other departments and is currently placed in a different building. It may also be asked whether these evaluations are taken seriously? Again, after analyzing internal reports and discussions, we feel they are. IEG creates a whole range of products that comment on the evolution of World Bank's projects and these reports are mostly based on the ES/EVM and PPAR data and the project reports behind their construction. Inside the World Bank there is a remarkable attention to the evolution of the ratings, so far that in 1992 after a long-lasting decline in project performance ratings, the Wapenhans Report (Wapenhans 1992) called for immediate changes in the operational structure of the Bank, as it "identified poor project monitoring and supervision, as well as a lack of focus on project outputs and development impact, as contributing to poor project portfolio performance" (Sigurdsson and Schweitzer 1995). Another question which the nature of this data may spark is whether these evaluations are reliable. This may present two types of concern. The first and more general one regards how we define successful a project. It is clearly a micro-definition, which heavily depends on the project; in fact its outcome variable measures the "extent to which the operation's major relevant objectives were achieved, or are expected to be achieved, efficiently"[19]. Therefore, a project is successful if it reaches the stated objectives. This may lead to think that project results may depend on how carefully their objectives are stated, for example a project may perform exceptionally well given extremely low goals or viceversa. The World Bank addresses this issue substantially in order to avoid such problems, by requiring an intense project review process within sectors, questioning whether the project is sufficiently ambitious, feasible and meets the poverty-fighting objectives guiding the World Bank mission. In order to attain comparability standards several departments are involved in a collective review of each project: the legal Vice Presidency, the procurement Vice Presidency, the regional Vice Presidency, the sector manager within the regional Vice Presidency, and for complex loans a whole committee, including the Network anchors[20] and other Vice Presidencies. A second concern may regard the credibility of these ratings, especially for what concerns the ES/EVM relative to PPAR. In their paper Denizer, Kaufman and Kraay (2011) do not find significant differences between the two evaluation types and their dummy capturing this difference is insignificant. In our study, we adopt different variables from their paper, still we do not find significant differences.

[19] Refer to
http://web.worldbank.org/external/default/main?theSitePK=1324361&pagePK=64253958&contentMDK=20789730&menuPK=5039271&piPK=64252979
http://web.worldbank.org/external/default/main?theSitePK=1324361&pagePK=64253958&contentMDK=20789685&menuPK=5039271&piPK=64252979

[20] By Network Anchor it is intended the "sectors" in which the Bank's work is developed, these are: Agriculture and Rural Development, Energy and Mining, Environment, Global Information/Communications Technology, Transport, Urban Development and Water, Economic Policy, Education, Financial and Private Sector Development, Financial Sector, Health Nutrition and Population, Poverty Reduction, Public Sector Governance, Social Development and Social Protection.

Annex C: The definitions of ICRG variables

The definition of democratic accountability is reported from the International Country Risk Guide: "This is a measure of how responsive government is to its people, on the basis that the less responsive it is, the more likely it is that the government will fall, peacefully in a democratic society, but possibly violently in a non-democratic one". The definition of democratic accountability is reported from the International Country Risk Guide: "This is a measure of how responsive government is to its people, on the basis that the less responsive it is, the more likely it is that the government will fall, peacefully in a democratic society, but possibly violently in a non-democratic one". The definition of government stability is reported from the International Country Risk Guide: "This is an assessment both of the government's ability to carry out its declared program(s), and its ability to stay in office. The risk rating assigned is the sum of three subcomponents, each with a maximum score of four points and a minimum score of 0 points. A score of 4 points equates to Very Low Risk and a score of 0 points to Very High Risk." The definition of bureaucracy is reported from the International Country Risk Guide: "The institutional strength and quality of the bureaucracy is another shock absorber that tends to minimize revisions of policy when governments change. Therefore, high points are given to countries where the bureaucracy has the strength and expertise to govern without drastic changes in policy or interruptions in government services. In these low-risk countries, the bureaucracy tends to be somewhat autonomous from political pressure and to have an established mechanism for recruitment and training. Countries that lack the cushioning effect of a strong bureaucracy receive low points because a change in government tends to be traumatic in terms of policy formulation and day-to-day administrative functions." The definition of corruption is reported from the International Country Risk Guide: "This is an assessment of corruption within the political system. Such corruption is a threat to foreign investment for several reasons: it distorts the economic and financial environment; it reduces the efficiency of government and business by enabling people to assume positions of power through patronage rather than ability; and, last but not least, introduces an inherent instability into the political process". The definition of ethnic tension is reported from the International Country Risk Guide: "This component is an assessment of the degree of tension within a country attributable to racial, nationality, or language divisions. Lower ratings are given to countries where racial and nationality tensions are high because opposing groups are intolerant and unwilling to compromise. Higher ratings are given to countries where tensions are minimal, even though such differences may still exist". The definition of external conflict is reported from the International Country Risk Guide: "The external conflict measure is an assessment both of the risk to the incumbent government from foreign action, ranging from non-violent external pressure (diplomatic pressures, withholding of aid, trade restrictions, territorial disputes, sanctions, etc) to violent external pressure (cross-border conflicts to all-out war)". These risk variables space over a given set, different between the variables, i.e. corruption ranges from 0 to 6, government stability from 0 to 12.